How can video analysis help laparoscopic surgeons?

P. Sánchez-González, I. Oropesa, S. Díaz,
Bioengineering and Telemedicine Centre (GBT), ETSI Telecomunicación. Universidad Politécnica de Madrid (UPM),
Madrid, Spain
Networking Research Center on Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN),
Madrid, Spain

psanchez@gbt.tfo.upm.es

F.M. Sánchez-Margallo

Minimally Invasive Surgery Centre Jesús Usón (CCMIJU).
Cáceres, Spain

P. Lamata, E.J. Gómez
Networking Research Center on Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN),
Madrid, Spain
Bioengineering and Telemedicine Centre (GBT), ETSI Telecomunicación. Universidad Politécnica de Madrid (UPM),
Madrid, Spain

Abstract—**Automatic analysis of minimally invasive surgical (MIS) video has the potential to drive new solutions that alleviate existing needs for safer surgeries: reproducible training programs, objective and transparent assessment systems and navigation tools to assist surgeons and improve patient safety. As an unobtrusive, always available source of information in the operating room (OR), this research proposes the use of surgical video for extracting useful information during surgical operations. Methodology proposed includes tools' tracking algorithm and 3D reconstruction of the surgical field. The motivation for these solutions is the augmentation of the laparoscopic view in order to provide orientation aids, optimal surgical path visualization, or preoperative virtual models overlay.**

Keywords: MIS, video image analysis, 3D reconstruction, tool tracking, organ reconstruction

I. Introduction

During an intervention, minimally invasive surgeons must navigate the anatomical landscape without the usual sensory clues. The deployment of this surgery is based on an indirect manipulation of specialized surgical tools introduced in the abdominal cavity through small incisions. Surgeons follow the operation through an indirect visualization of the surgical field captured by an endoscope [1][2].

Surgical navigation systems are providing new possibilities for minimally invasive surgery (MIS). These systems allow to transfer preoperative data, images and decisions to the operating room (OR), and to give the surgeon guidance during the procedure. Traditionally, the surgeon had to translate all this information mentally to the real scenario in front of him, with the difficulties this entails, most notably disorientation with respect to the medical image studies. Image guided surgery (IGS) systems have the potential to compensate these limitations. Difficulties arise in soft-tissue surgeries due to organ shifting and tissue deformation, caused by breathing, heartbeat, patient movement, and surgeon's manipulation [3].

Automatic analysis of surgical video captured by the endoscope has the potential to drive new solutions that lighten existing needs for safer surgeries. Reproducible training programs, objective and transparent assessment systems and navigation tools to assist surgeons and improve patient safety can be developed by means of the analysis of video sequences.

Surgical video sequences provide useful information about the position of instruments and organs, surgical maneuvers, measurements of distances or even an approximate 3D reconstruction of the surgical scene. Laparoscopic video images are an always available source of information and can be used without adding extra technological components in the OR. This research proposes its use for extracting useful information during interventions, helping surgeons within their performance [4].

II. Methodology

A. Tools' tracking algorithm

During an intervention, keeping surgical tools centered in the field of view of the image is most important in order to reduce risks of damaging non visible anatomic structures. 3D localization of surgical instruments is also a critical issue and can help overcome the lack of depth perception and misorientation characteristic of MIS techniques. In these procedures, the knowledge of the tools' 3D position constitutes an important step to guide and advice surgeon about the proximity of delicate areas.

Traditionally, object tracking is achieved by devices based on mechanical, optical, acoustic or electromagnetic technologies [5]. Their introduction in the OR may compromise patient safety since they disturb the surgical workflow and ergonomics.

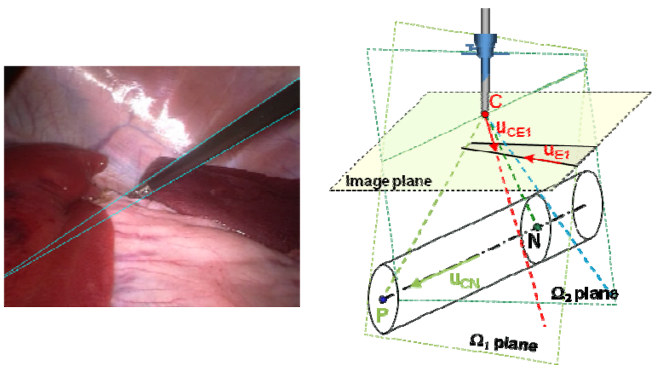
Video analysis can cope with 2D and 3D tools' localization. Tools' segmentation has been addressed using several approaches, using color either on RGB [6][7] or HSV ("Hue,

This research has been partially funded by the THEMIS project of Networking Research Center on Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN) and the TELMA project (TSI-020110-2009-85) of the Spanish Ministry of Industry, Tourism and Trade.

Saturation, Value") space [8], geometrical properties such as border detection [9], or a combination of both [10].

Moreover, tracking methods must cope with partial occlusions, the insertion and removal of different tools, and with image quality degradation caused by gas and smoke.

Image-guided surgical applications require real-time, accurate and robust tracking of tools during the entire surgical procedure. The challenge of the approach is to extract the 3D position and orientation of a rigid cylindrical tool from the 2D information of the surgical scenario captured by endoscope. The method proposed extracts information about the tool's edges and tip employing automatic segmentation based on color and movement features, and based on the projective geometry of the surgical scene (figure 1.b) [11] is able to determine its 3D position and orientation.



(a) (b)

Figure 1. (a) Image processing for extracting 2D information (b) Model of the laparoscopic tool used for calculating the 3D position and orientation of the tools; C: optical center, P: tip of the tool; N: point of the cylinder axis whose projective line is perpendicular to this axis; Ω1, Ω2: planes of sight of the tool edges; E1, E2: projective tool edges.

B. *Surgical scene reconstruction methods*

One of the limitations in laparoscopic surgery is the loss of depth information. The recovery of the depth map associated to the image from the endoscopic camera is the objective of this research. There are in the literature several computer vision approaches to the problem [11]-[18], but with limited success. The main challenge is the complexity of the surgical scene, which leads to the limitations of existing solutions. Most techniques assume some physical constraints not always valid for an endoscopic scene (e.g. a static scene viewed from different angles by a moving camera), and occlusions or camouflages (occult zones or not detected ones) are difficult to resolve.

In order to achieve a reconstruction of the surgical scene, this research exploits one of the depth clues in an image: the shading of objects [19][20]. "Shape from Shading" (SfS) algorithms are applied, which assume a lambertian light interaction model with the surface of organs. The strength of this approach is the only need of one frame of a video sequence to estimate a depth map, and the main limitation is the impossibility to handle occlusions and relative depths between adjacent objects. With these characteristics, one of the natural applications of SfS solutions is laparoscopic liver procedures. The goal is integrating visualization of major vessels and targeted tumors as well as preoperative resection planning information into the endoscopic images [20].

In more detail (see Figure 2), the approach developed for liver interventions implements a propagation SfS technique initialized from the closest point to the camera. Due to the specular reflection of light on the surface of the liver, the closest point in the image is the brightest. This solution therefore reconstructs the visible patch of the liver captured by the endoscope, with a previous segmentation step as described in [4]. The reconstruction problem has two unknown properties of the object, the distance to the camera (what is coupled to knowing the scale of the object), and the albedo of the material (the property that determines how much light is absorbed by the material, and how much is reflected back). The solution to these two unknowns is resolved by an iterative 3D reconstruction and 3D registration via a modified Iterative Closest Point algorithm (ICP). For more details, the reader is referred to [20].

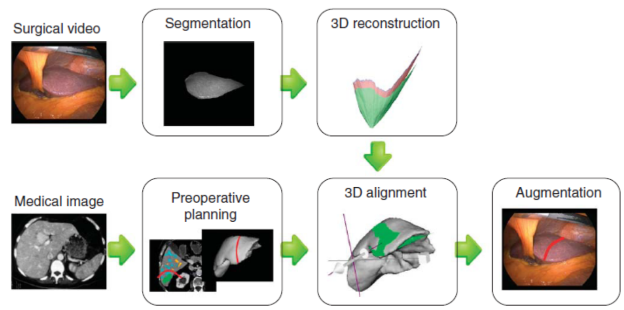


Figure 2. Data pipeline for the augmentation of the surgical view in a liver resection procedure, where the resection line decided in a preoperative stage is overlaid on the surgical image.

III. RESULTS

A. *Tools' tracking results*

Two validation sequences are used to evaluate the performance of the laparoscopic tool's tracking method. Two different movements were simulated, "Constant depth movement" and "Variable depth movement", acquired in a laboratory surgical setting. With this setting results obtained showed a standard deviation lower than 2.09 and 3.55 in constant depth movements and variable depth ones. However, there is an inherent constant error (lower than 9 mm. in variable depth movements), whose origin is due to inaccuracies in the tool's edges detection.

B. *Surgical scene reconstruction results*

An exemplary result of the 3D reconstruction of a surgical scene is presented in Figure 3, combined with the estimation of the 3D pose of the surgical tools. An important limitation of shading information is the presence of noise sources, like inter-reflections or specularities. While specular bright regions are relatively easy to detect and compensate, inter-reflections originated in surrounding organs are quite unpredictable and variable due to movements caused by respiration, cardiac motion, and mechanical shift accompanying the surgical intervention.

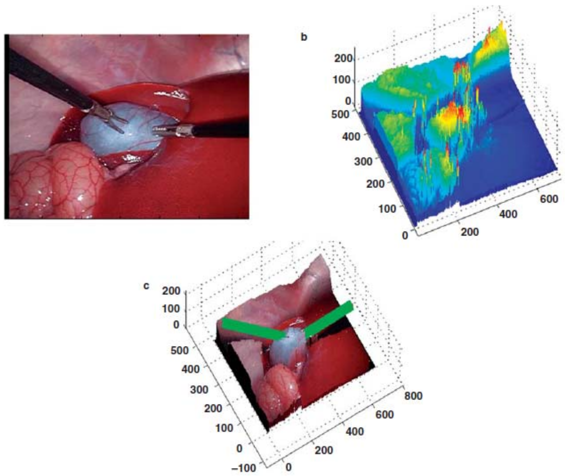


Figure 3. Reconstruction of the surgical scene. (a) Original laparoscopic image, (b) depth map, (c) image 3D reconstruction combining scene reconstruction and tool localization (green cylinders).

IV. CLINICAL APPLICATIONS

A. Surgical guidance and orientation

In order to deal with deviations from preoperative data due to organ shifting and deformations, updated information must be given periodically to update the navigational models. Typically, new images based on intraoperative CT or laparoscopic ultrasound (LUS) are taken. However, inclusion of these systems in the OR contributes negatively to a technological overload of elements. Plus, their usage constitutes an interruption in the normal workflow of the intervention. Additionally, tool's tracking for guidance is usually achieved by mechanical, optical, acoustic or electromagnetic sensors [5].

By means of endoscopic scene reconstruction, updated information on the organ's disposition can be acquired and registered with the preoperative models in a non-disruptive fashion. Similarly, tools can be tracked with respect to the endoscope without the need of mounted sensors. The surgical workflow needs not be seriously interrupted, and no additional equipment is necessary in the OR (save for visualization),

B. Objetive motor skills assessment

Metrics definition extracted from laparoscopic video images can cope with efficiency kinematic parameters, such as path length, path deviation, speed profile or motion smoothness [21].

Surgical video recordings can also be used for objective motor skills assessment [22] and for post-surgical evaluation of the operation. The latter can be achieved by means of the automatic extraction and storage of "quality parameters" of the surgical intervention in the patient's Health Care Record. Once an intervention is finished, a clinical report can be performed with data extracted from the automatically processed video.

V. DISCUSSION

Image guided assistance and navigation for laparoscopic surgery poses as an important feature for overcoming some of the most problematic limitations a surgeon must deal with. However, development of these systems has been slow due to the complexities derived from soft tissue deformations and the inexistence of anatomical landmarks for patient-to-preoperative model registration.

The traditional approach, inherited from neurosurgical navigators, has often opted for a combination of three elements: (1) active sensing for both camera and tools to allow guidance, (2) intra-operative image modalities such as C-arms or LUS for updated information of anatomical structures' position(3) the inclusion of artificial fiducials that act as made-shift landmarks for organ registration. As has been pointed out , technological overload and surgical workflow derived from this approach may become a handicap for patient surgery.

Laparoscopic video, an ever-present resource in MIS techniques, may help overcome some of these limitations. Applied computer vision for extraction of useful information of the surgical scene can serve as an aid to surgeons, in a transparent, non intrusive manner.

Surgical tools' localization can provide useful information for image-guidance on intra-operative procedures in MIS. In these techniques, the knowledge of the tools' 3D position constitutes an important step to guide and advice surgeon about the proximity of delicate areas. This information can be completed with the 3D reconstruction of the organs and the anatomic structures present in the scene.

Passive tracking of the surgical tools is possible without the use of sensor systems. However, pinpoint accuracy is required in these scenarios. Moreover, a compromise must be found between system robustness and real-time updates. Image processing tends to be complicated, due to the high variability in the intraoperative feeds. Thus, in order to ease processing times while ensuring acceptable results, the use of colour markers may be required on the surgical instruments.

A limiting factor in image-based tracking of tools is the presence of blurred tool’s edges caused by its movement. This causes inaccurate edge detection, which can then propagate to the 3D pose estimation. A robust edge detection can include temporal constrains, an also geometrical constrains based on the position and projection of the insertion point (trocar point) [6], which can be obtained with the 3D orientation information extracted from the proposed method.

Complementary to tools' tracking is the 3D reconstruction of the surgical scene. Updating pre-operative models based on endoscopic information poses itself as the alternative to having image acquisition systems in the OR. However, the reconstruction of accurate surgical models in real time must be first addressed. Moreover, video information is constrained to the field of view of the endoscope, and thus model update can be only partially and superficially achieved. Taking this into account, the use of non-invasive LUS as a supporting technology can be considered for complete registration.

Characteristics of the reconstruction of a 3D model from laparoscopic images will be strongly affected by its application and its desired accuracy, robustness, velocity or computing cost will be constrained by this issue. Thus, the handling of laparoscopic image data in IGS requires real-time algorithms and also careful synchronization; in particular between simultaneously acquired data from heterogeneous sources (the registration problem). Accuracy and robustness needed in this case scenario are extremely important, because of risks carried out when operating a patient in the OR. Moreover, accurate methods have to be developed to register image space to physical or patient space based on video processing, with the main problem being the lack of anatomical landmarks. In this registration field, there are important challenges to be tackled such as adaptation to the variability of quality and appearance of images, strongly affected by surgical actions and also by inter-patient biological variability.

All limitations considered, including camera-related aspects such us image quality, distortions or the field of view, it is likely that video-based navigators will probably complement, and not replace, other existing solutions. It is our belief that our main key innovation is to provide navigation systems based on both medical images and computer vision analysis of the surgical video, without introducing tracking hardware in the OR nor modifying the surgical workflow, developing new usable workstations which will be easily integrated in the operating theatre.

Research efforts are required to further develop these technologies, so as to harness all the valuable information available in any video-based surgery. Future research will deal with the optimization of robustness, accuracy and processing time of the algorithms in order to make easier their incorporation in the clinical routine.

VI. CONCLUSIONS

Automatic analysis of MIS video has the potential to drive new solutions for alleviating needs of development new navigation systems to assist surgeons during the operating techniques and improvement patient safety. Thus, video-endoscopic image analysis is essential for extraction of relevant information of the surgical workspace and for creating real-time image-guided surgical aids: tools', organs' and endoscope tracking can be achieved by means of its analysis.

REFERENCES

[1] Usón J, Sánchez-Margallo FM, Pascual S, Climent S. Formación en Cirugía Laparoscópica Paso a Paso. 4th ed. Minimally Invasive Surgery Centre Jesús Usón eds. Cáceres, Spain; 2010.

[2] Cuschieri A. Laparoscopic surgery: current status, issues and future developments. Surgeon; 3: 125-130 . 2005.

[3] M. Baumhauer, M. Feuerstein, H.P. Meinzer, J. Rassweiler. "Navigation in endoscopic soft tissue surgery: perspectives and limitations". J Endourol. 22(4), pp. 751-66. 2008.

[4] P. Sánchez-González, A.M. Cano, I. Oropesa, F.M. Sánchez-Margallo, F. del Pozo, P. Lamata, E.J. Gómez. "Laparoscopic Video Analysis for Training and Image Guided Surgery". Minimally Invasive Theraphy & Allied Technologies MITAT (Epub ahead of print). 2010.

[5] M.K. Chmarra, C.A. Grimbergen, J. Dankelman. "Systems for tracking minimally invasive surgical instruments". Minimally Invasive Therapy and Allied Technologies, 16 (6), pp. 328 – 340. 2007.

[6] S. Voros, J.A. Long, P. Cinquin. "Automatic localization of laparoscopic instruments for the visual servoing of an endoscopic camera holder". Med. Image Comput. Comput. Assist. Interv., 9(1), pp. 535-542. 2006.

[7] A. Casals, J. Amat, E. Laporte. "Automatic guidance of an assistant robot in laparoscopic surgery". Proceedings 1996 IEEE International Conference on Robotics and Automation, 1, pp. 895-900. 1996.

[8] A. Krupa, J. Gangloff, C. Doignon, et al. "Autonomous 3-D positioning of surgical instruments in robotized laparoscopic surgery using visual servoing". IEEE Transactions On Robotics And Automation - Special Issue On Medical Robotics, 19 (5), pp. 842-853. 2003.

[9] O. Tonet, T.U. Ramesh, G. Megali, P. Dario. "Tracking endoscopic instruments without localizer: image analysis-based approach". Stud. Health Technol. Inform., 119, pp. 544-549. 2006.

[10] S.J. McKenna, H. Nait Charif, T. Frank. "Towards Video Understanding of Laparoscopic Surgery: Instrument Tracking". Conf. Image and Vision Computing New Zealand 2005. pp. 317-321. 2005.

[11] A.M. Cano, F. Gayá, P. Lamata, P. Sánchez-González, E.J. Gómez. "Laparoscopic tool tracking method for augmented reality surgical applications". Lecture Notes in Computer Science (LNCS), 5104, pp. 191–196. 2008.

[12] R. Zhang, P.S. Tsai., J.E. Cryer, M. Shah. "Shape from Shading: A Survey". IEEE Transactions on pattern analysis and machine intelligence, 21, pp. 690-706. 1999.

[13] Samaras D., Metaxas D., Fua P., Leclerc Y.G. Variable Albedo Surface Reconstruction from Stereo and Shape from Shading. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'00) – 2000, 1: 1480.

[14] J. Banks.,M. Bennamoun, K. Kubik, P. Corke. "An accurate and reliable stereo matching algorithm incorporating the rank constraint". Symposium on Intelligent robotic systems, pp. 23-32. 1999.

[15] P.M. Newman. "On the structure and solution of the simultaneous localization and map building problem". Ph.D. dissertation, University of Sydney, 1999.

[16] T. Jebara, A. Azarbayejani, A. Pentland. "3D Structure from 2D Motion". IEEE Signal Processing Magazine, 16 (3), pp. 66-84. 1999.

[17] F. Dellaert, S. Seitz, C. Thorpe, A. Thrun. "Structure from Motion without Correspondence". IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'00), 2000.

[18] A. J. Davison, I. Reid, N. Molton, O. Stasse. "Real-time single camera SLAM". IEEE Trans. Pattern Anal.Mach. Intell., 29(6), pp. 1052–1067. 2007.

[19] P. Sánchez-González, F. Gayá, A.M. Cano, E.J. Gómez. "Segmentation and 3D reconstruction approaches for the design of laparoscopic augmented reality environments". Lecture Notes in Computer Science (LNCS) 5104, pp. 127–134. 2008.

[20] P. Lamata, T. Morvan, M. Reimers, E. Samset, J. Declerck. "Addressing Shading-Based Laparoscopic Registration". World Congress on Medical Physics and Biomedical Engineering (WC2009), pp. 189–92. 2009.

[21] R.M. Satava, A. Cuschieri A, Hamdorf J. Metrics for objective Assessment. Surg Endosc, 17, pp. 220-226. 2003.

[22] I. Oropesa, P. Lamata, P. Sánchez-González, J.B. Pagador, M.E. García, F.M. Sánchez-Margallo, E.J. Gómez. Virtual Reality Simulators for Objective Evaluation on Laparoscopic Surgery: Current Trends and Benefits. In J. Kim eds. Virtual Reality. Rijeka: Intech Open Access Publisher, pp. 349-374. 2011.